New tools for use in the musicology of record production

Kirk McNally, George Tzanetakis, Steven R. Ness
University of Victoria
Music and Computer Science
kmcnally@uvic.ca, gtzan@cs.uvic.ca, sness@sness.net

Abstract

This paper introduces a stereo 3-D panning visualization tool based on methods borrowed from the field of Music Information Retrieval (MIR). This tool helps to illustrate and quantify production decisions and recording practices used by engineers and producers in the record production process. The tool is also valuable for pedagogical purposes, providing students with a visual feedback of what they are (or are not) hearing in recordings as they develop their critical listening skills. A case study comparing a body of work by Tchad Blake and Rick Rubin illustrates the value of this tool to the musicology of record production.

1. Introduction

The field of Music Information Retrieval (MIR) is defined as, "a multi-disciplinary research endeavor that strives to develop innovative content-based searching schemes, novel interfaces, and evolving networked delivery mechanisms in an effort to make the world's vast store of music accessible to all" [1]. How this relates to the musicology of record production may at first appear a tenuous link. Consider however that in an effort to better understand the music being searched, MIR researchers seek to develop tools that examine and extract features of the music. A system might evaluate the timbral similarities between different styles of music, eg. rap vs. jazz, and produce genre classifications based on these results. Expand this thinking to incorporate all of the different technical treatments and techniques at the disposal of the recording engineer and producer and, with accurate and robust enough tools, it maybe possible to not simply classify genres, styles or tempos, but also the engineer and producers individual stylistic techniques or "signature". Just as the studying of scores by young composers provides insight into the masters' techniques and trademarks, the 3-D panning visualization tool introduced in this paper allows student engineers and producers insight into the technical attributes that makes an album distinct, ie. a "Rick Rubin" album vs. a "Phil Spector" album.

1.1 Justification

Prior work with colleagues in the field of MIR shows that including stereo panning features to classify record production style improves scores for genre classification of a music database versus "classic" methods based solely on timbral similarities for genre classification. Classification accuracies improved in the order of

ten percent for the given task of distinguishing 1960's garage music from 1980-90's grunge music. An increase of twenty percent was seen for the task of distinguishing acoustic vs. electric jazz [2]. While this may seem a trivial task to the audio engineer or producer (indeed it is hypothesized that a trained listener would score higher than the MIR system) it should not be discounted. The work clearly showed that there are panning decisions made as part of the production process that can be identified as "stylistic" or "accepted" for a given genre. It is argued that this ----part of the production process is natural, indeed, instinctive for an industry professional – it just sounds right! Again, expand this concept to an individual producer or engineer with years of experience, one who has evolved to have a so-called "sound" and classification of this individual style or "genre" maybe possible. This work also identified that the current means for visualizing this panning information was poor and non-intuitive. Developing a new method for visualizing this information, one that would enable it to be more accessible to scholars, was the initial goal of this project.

1.2 Panning Pedagogy

In designing and developing the visualization tool it became clear that a valuable pedagogical tool was also being created. Students of audio engineering and record production work to develop their listening skills through critical listening exercises and practical work in the studio. The process of mixing (the place where most panning decisions are made) is a complex problem of balancing levels vs. panning vs. frequency vs. effects. There is an established history of thinking about this mixing process visually, where individual instruments are drawn by students to create a "picture" of their final mix [3]. The analysis of commercial recordings is the counterpoint to this and will guide students in their own mixing decisions. What is missing from this methodology is that while students are developing their technical ear they simply may not hear all the elements in a given mix. The tool introduced provides visual feedback of these elements to aid students in the development of their critical listening skills.

2. The Tool

2.1 Basic Design

The stereo 3-D panning visualization tool allows users to visualize the panning of different elements in a recording, with panning graphed on the x-axis, frequency on the y-axis, and the real-time component creating a waterfall-type display on the z-axis. The tool computes the panning index, a frequency-domain source identification system based on a cross-channel metric as described in Avendano [4]. This panning index, giving values of -1.0 to 1.0, full left to full right respectively, allows for the mapping of the individual frequency components in different FFT (Fast Fourier Transfrom) or MFCC (Mel Frequency Cepstrum Coefficient) bins along the left-right panning dimension. As each FFT of MFCC bin is computed it is then mapped to zero time on the z-axis. As each new bin is

calculated the previous bin moves along the z-axis (away from the user), thus creating a "picture" of the track being visualized. Magnitude (level) of the individual frequency components are mapped using a colour scale from green (low level) to red (high level). This should be familiar to users of and digital audio workstation (DAW). Using the opening bars of The Beatles *In My Life* to illustrate this*,* we see the opening guitar and bass lines panned fully to the left (panning index = -1.0). John Lennon's vocals then enter panned fully right (panning index +1.0) and Ringo Star's drums enter again panned left (panning index -1.0).

An important side effect of this plotting process is that the ambience or reverb present in the given track is also clearly visualized. If the sound source is a "dry" amplitude panned source, such as the opening lines of the Rubin produced Jay Z track *99 Problems*, we see a single line in the center of the display. Conversely, if the source is panned to the center but accompanied by reverb, as in the Blake produced track, *Name*, by Artist, we see a wider image that "spreads" between left and right on the display.

2.2 Software Design

Two versions of the software were developed, the first, a web-based version of the tool using a Flash application that was embedded within a standard HTML web page. Due to the recent addition of native 3D support to Flash 10, it was possible to achieve near real-time performance in displaying the 3D panning information for a given track.

Marsyas (http://marsyas.sf.net) a Music Information Retrieval framework was used to generate files containing data points that represented the 3D panning information for a given track. This version of the tool has the advantage of being easily accessed by users from around the world using a wide variety of computers and operating systems. However, in order to accurately visualize the transient information of a given track large amounts of data need to be drawn on the screen in a very short period of time. Even with the 3D support in it's latest version, Flash is unable to display the requisite number of data points in real-time yielding a display lacking the resolution required. For this reason, a second version of the software was developed, again using the Marsyas programming framework. This version was combined with hardware accelerated 3D using the OpenGL toolkit and a GUI created using the Qt framework. This application takes advantage of the hardware 3D graphics cards present in all modern computer systems to display large numbers of data points simultaneously, and as the entire application is written in C++, it is able to provide real-time performance in many use cases. This tool is open source, free software and is available for download at the main Marsyas website (http://marsyas.sf.net).

3. A case study: Rick Rubin vs. Tchad Blake

Two established industry professionals were selected for a case study with the goal of clearly identifying an individual style or

technique with regards to panning in the production process.

Rick Rubin is a Grammy award-winning producer successful in a variety of genres, which provides a body of work well suited to test the hypothesis that a "signature" exists in the albums he produces. Rubin's productions are known to exhibit a stripped-down sound, one that eschews the use of reverb, instead using naked vocals and bare instrumentation. Rubin is quoted in a NYTimes article saying, "There's just a natural human element to a great song that feels immediately satisfying. I like the song to create the mood" [5]. Though he describes himself as, "no expert at the technical aspects" of the record production process [6] this underlying production goal of allowing the song to, "create the mood" would certainly seem to dictate albums that are very sparse and natural. The selected works for Rubin include: Beastie Boys *Liscenced to Ill,* Jay-Z *The Black Album,* Dixie Chicks *Taking the Long Way,* Red Hot Chili Peppers *Blood Sugar Sex Magik,* Tom Petty *Wildflowers* and Johny Cash *American IV: The Man Comes Around.*

Tchad Blake is an engineer who describes himself as, "...really an engineer/producer, not a composer/ producer...I like working with artists who have strengths in arrangements and are musically adept. I think I am good at contributing atmosphere and helping them flesh things out". [7]. When asked his ideology and approach to panning on the *Gearslutz* forum Blake responds, "Anywhere and everywhere. Mostly hard. Outstide the speakers if I can" [8]. Blake is also known for his use of binaural techniques and an affinity towards dynamic mixing textures with unique sonic textures. The selected works for Blake include: Sheryl Crow *The Globe Sessions,* Joseph Arthur *Redemption Son*, Pearl Jam *Binaural,* Elvis Costello *Brutal Youth* and Suzanne Vega *Beauty and Crime.*

3.1 Methodology

A total of seventy-two Blake produced/engineered tracks and seventy-four Rubin produced tracks were used for the case study. Observations were first made using the visualization tool. It was found that the Rubin produced tracks exhibit limited dynamic panning and are primarily driven by very dry, mono vocals. Panning of mono elements in the mixes tended towards extreme L/C/R panning. The stereo elements eg. drums or piano, generally were seen to be spread full left to right, again with little or no ambience or reverb associated with them. The Blake tracks show more instances of dynamic panning and a much greater use of reverb and effects. Blake's vocals are often presented with reverb or effects, but even when presented without reverb, this presentation will inevitably change. For example in a chorus or bridge with the vocal being augmented by reverb or effects. This classic "lifting" of the chorus is also present in Rubin produced tracks but his technique uses track arrangements rather than the use of effects. A final observation was in the way that Blake and Rubin treated low frequency information (below 500Hz) with regards to reverb. Blake has a

tendency to spread information below 500Hz through the use of reverb or ambiences. In the instances where Rubin does use reverb it doesn't generally extend into this low frequency region. This creates a "thicker", although less dynamic sound for Blake's tracks in this frequency range with Rubin's sounding more precise and dynamic.

3.2 Machine Learning Classification

In order to support these observations and provide empirical data showing them to be true, the selected tracks were run through a MIR machine learning classifier. The basic analysis tool is the same as that previously described, computing the panning index for individual frequency components of FFT or MFCC bins. For this component of the study the process of plotting the data for visualization has simply been replaced by computational analysis. In order to compare equal segments of audio the tracks were first segmented into clips of 30 seconds each. Any end sections of the songs that were less than 30 seconds in length were removed. The panning information, or "features", were then extracted from each of the clips using Marsyas. The features extracted were the stereo panning index for low (0Hz-250Hz), middle (250Hz-2500Hz) and high (2500Hz-22050Hz) frequency bands as outlined in Tzanetakis [1] and Avendano [2]. The output of this process provided the mean and standard deviation of these audio features calculated over a number of different window sizes. These audio feature vectors were then used to train a Support Vector Machine (SVM)

classifier. This trained machine learning classifier was then used to classify the tracks using a standard 10-fold cross validation methodology. The results of this process showed that 70.6 percent of the time the tracks were classified correctly for the given task of identifying Blake vs. Rubin.

4. Conclusions

The case study shows that the observations made using the visualization tool can be trusted and that there is a difference in the way the two producer/engineers, Rubin and Blake, treat panning in the record production process. It must be acknowledged that panning alone is quite a crude measure of production style. However, any result showing that record production can be quantified must be seen as positive for this area of study. If the process can be quantified then it can be separated from the artist, thus the producer/engineers role and impact on the finished track can be clearly studied and appreciated.

Readers interested in this tool should contact Steven Ness for download instructions at sness@sness.net. They can also obtain the MarPanning source code by downloading Marsyas from http://marsyas.sf.net

4. Future Work

The primary focus of future work will be to create a VST plug-in version of the tool that can be used in any DAW. This would allow users to visualize the track throughout the recording and mixing process to help guide their work. Another goal is to

increase the number of features for the classification of record production style. Ideas include tools to quantify the use of compression, delay and other effects, and to deconstruct track arrangements through sound-source separation techniques. With an increased number of features an even greater ability to classify individual production techniques will be possible.

5. Acknowledgements

I would like to acknowledge Randy Jones for the ideas and discussions that led to this work, Peter Driessen for his valuable input, Steven Ness for his programming and finally George Tzanetakis for his expertise in MIR.

6. Bibiliography

[1] Tzanetakis, G., Jones, R. and McNally, K. Stereo Panning Features for Classifying Recording Production Style. In Proceedings of the International Conference on Music Information Retrieval (ISMIR), Vienna, Austria, September 2007.

 [2] Avendano, C. "Frequency-domain source identifica-tion and manipulation in stereo mixes for enhance-ment, suppression and re-panning applications. In Proceedings of IEEE Workshop on Applications of Signal Pro-cessing to Audio and Acoustics (WASPAA), pages 55– 58, 2003.

[3] J. Stephen Downie. Toward the Scientific Evaluation of Music Information Retrieval Systems. In Holger H. Hoos and David Bainbridge Proceedings of the Fourth International Conference on Music Information Retrieval: ISMIR 2003. (pp. 25-32).

[4] Eargle, John. Handbook of Recording Engineering 3rd ed. New York: Van Nostrand Reinhold Co. 1986.

[5] Hirschberg, Lynn. "The Music Man". *The New York Times* (September 2007). <http://www.nytimes.com/2007/09/02/magazine/02rubin.t.html?_r=1> (accessed 30 October 2009).

[6] Tyrangiel, Josh. "Rick Rubin: Hit Man". *Time Magazine* (February 2007)*. <http://www.time.com/time/magazine/article/0,9171,1587248,00.html> (accessed 30 October 2009).

[7] Bonzai, Mr. "Working in the Real World: Engineer/Producer Tchad Blake". *Mix Magazine* (February 2005). <http://mixonline.com/mag/audio_engineerproducer_tchad_blake/index.html> (accessed 30 October 2009).

[8] Q&A with Tchad Blak*e. <http://www.gearslutz.com/board/q-tchad-blake/121086-tchad-blake-resume.html> (accessed 30 October 2009).